# iassist

### international association for social science information service and technology

### l'association internationale pour les services et techniques d'information en sciences sociales

# NEWSLETTER
## vol. 1, no. 3
## MAY 1977

### TABLE OF CONTENTS

The IASSIST Newsletter represents an international cooperative effort on the part of individuals managing, operating, or utilizing machine-readable data archives, data libraries and data services. The Newsletter has as its primary intent the dissemination and exchange of information on significant developments in these information centers. Four times a year this Newsletter will report on activities related to the production, acquisition, preservation, processing, distribution, and utilization of machine-readable data carried out by its members and others in the international social science community. Your contributions and suggestions for topics of interest are encouraged and welcomed.

The Newsletter is published in November, February, May, and August. Articles, reports, announcements, and other notes for inclusion in the Newsletter must be received by the Editor no later than October 31st, January 31st, April 31st, and July 31st. If a delay in a report is anticipated, the writer should alert the Editor either by telephone or letter before the above cited dates. The Editor will attempt to delay publication by no more than 7 days.

Information for inclusion in the Newsletter should be typewritten and doublespaced. Authorship of all information should be clearly identified (e.g., "submitted by ... [full name, address, and telephone number]"). Action Group reports should include the names of the participants and their national and institutional affiliations. If an individual(s) chaired a meeting, this information should be noted. The format established in the first volume of the Newsletter should be followed when preparing the Action Group report.

# iassist NEWSLETTER vol. 1, no. 3

## IASSIST STEERING COMMITTEE

Carolyn Geda, CHAIRPERSON, Inter-university Consortium for Political and Social Research, P. O. Box 1248, University of Michigan, Ann Arbor, Michigan 48106 USA

Ed Hanis, TREASURER, Social Science Computing Laboratory, University of Western Ontario, London, Ontario, Canada

Manuel Mora y Araujo, SOUTH AND CENTRAL AMERICAN SECRETARIAT, Centro de Investigaciones Sociales, Instituto Torcuato Di Tella, 11 de Septembre 2139 1428 Buenos Aires, Argentina

Sharon Chapple Henry, CANADIAN SECRETARIAT, Data Clearing House for the Social Sciences, 151 Slater, Ottawa, Ontario K1P 5N1 Canada

Naresh Nijhawan, ASIAN SECRETARIAT, Indian Council of Social Science Research, Data Archive, 35, Ferozshah Road, New Delhi-11001, India

Per Nielsen, WEST EUROPEAN SECRETARIAT, Dansk Data Arkiv, H.C. Andersens Boulevard 38 mezz., DK-1553 København V., Denmark

Krzysztof Ostrowski, EAST EUROPEAN SECRETARIAT, Komitet Nauk Politycznych, PAN, Sekcjz Przetwarzania i Archiwizacji Danych, skrytka pocztowa 12, 00-955 Warszawa

Judith Rowe, UNITED STATES SECRETARIAT, Computer Center, Princeton University, 87 Prospect Avenue, Princeton, New Jersey 08540 USA

Joseph Bonmariage, Belgian Archives for the Social Sciences, University of Louvain, SH-2, 1348 Louvain-la-Neuve, Belgium

Evert Brouwer, Amsterdam Social Science Data Archive, University of Amsterdam, Roeterstraat 15, Amsterdam, Netherlands

Ivor Crewe, Social Science Research Council Survey Archive, University of Essex, Wivenhoe Park, P. O. Box 23, Colchester, Essex CO4 3SO, England

Bjorn Henrichsen, Norwegian Social Sciences Data Services, University of Bergen, Christiesgate 15-19, N-5014 Bergen, Norway

Thomas Madron, Academic Computing and Research Services, 243 Grise Hall, Western Kentucky University, Bowling Green, Kentucky 42101 USA

Guido Martinotti, Archivio dati e programmi per le scienze sociali, via G. Cantoni 4, 20144 Milan, Italy

Cees Middendorp, Steinmetzarchief, Kleine-Gartmanplantsoen 10, Amsterdam-C., Netherlands

Ekkehard Mochmann, Zentralarchiv für Empirische Sozialforschung, University of Cologne, Bachemer Strasse 40, 5 Koln 41, Federal Republic of Germany

David Nasatir, Behavioral Sciences Graduate Program, California State College, Dominguez Hills, California 90747 USA

Alice Robbin, Data and Program Library Service, University of Wisconsin, 4452 Social Science Building, Madison, Wisconsin 53706 USA

## CHAIRPERSON'S REPORT

The second North American meeting was held in Toronto, Canada, May 11-12th, 1977. Sharon Henry, the Canadian Secretariat was able to secure funds from Canada Council to support the travel of participants delivering formal presentations at the opening session. Sixty-five individuals attended the conference. Abstracts of the presentations were submitted prior to the conference and translated to French. The English and French versions of the Abstracts, along with the complete English text of the presentations were available at registration. Participants of the conference will automatically receive a copy of the Proceedings and members of IASSIST will be able to purchase copies at a reduced rate. The registration fee for the conference will be used to defray reproduction costs of the Proceedings. Any additional monies from fees will be forwarded to the treasurer, Ed Hanis. The remainder of the meeting was devoted to the Action Group workshops. Abstracts of the presentations and the progress reports resulting from the Action Group workshops are published in the body of this Newsletter.

Per Nielsen, the West European Secretariat, has organized a European IASSIST Action Group workshop to be held at the Danish Data Archives in Copenhagen, June 26-29, 1977. The European AGs will build on the products of the US and Canadian efforts as well as previous work carried out by European archives, particularly in the area of classification and documentation. Proceedings of this meeting will be published. [See Newsletter section on the Danish Data Archives sponsored IASSIST West European meeting.]

The February 1978 IASSIST meeting has been scheduled for the 8th-12th, at Nordic Hills, Itasca, Illinois. The program is detailed in the Newsletter section describing upcoming meetings. The Chicago area was selected as the site of the meetings to permit maximum participation, given its central location in North America. Topics for presentation at the symposia are invited. Please submit the topics to Tony Falsetto, Public Readable Archives, Public Archives of Canada, 2850 Cedarwood Dr., Apt. 909, Ottawa, Ontario, Canada, with a copy to Carolyn Geda.

The IASSIST program for the meetings to be held in conjunction with the International Sociological Association in August 1978, has been approved. The schedule of activities is presented in the Newsletter section describing upcoming meetings.

Initial monitoring of IASSIST receipts and expenditures verifies that there are adequate funds to cover the cost of production of the Newsletter, mailing costs, and subscriptions to S S Data. Registration fees for the conferences are sufficient to support the costs of the meeting and publication of the Proceedings. A full treasurer's report will be included in the next issue of the Newsletter.

As has been noted in previous Newsletters, the working manual for cataloguing MRDF, compiled by Sue Dodd, University of North Carolina, will be the first IASSIST related publication. In recognition of this work and forthcoming pieces, the Steering Committee has proposed a set of publishing policies and has appointed an Editorial Board to work with authors and publishing houses in whatever capacity is facilitative. The Editorial Board is composed of Elina Almasy, Carolyn Geda, Edward Hanis, Cees Middendorp, and David Nasitir.

## SECRETARIAT REPORTS

United States Secretariat Report

Judith Rowe
Computing Center
Princeton University

The US region of IASSIST now has 78 paid members, from a total of 50 data ser- vice organizations. It was the intention of the Steering Committee to define IASSIST as a professional organization of individuals, rather than as an organiza- tion of institutional representatives, and therefore it is particularly gratifying to see multiple memberships from a number of data archives and data libraries. This is a trend which we would like to see continue, since it provides an opportunity for an organization to enjoy the benefits of participation in multiple Action Groups. Those members already involved in Action Group projects find it difficult to make substantial commitments to other action groups, although they recognize their per- sonal interest in and their organization's needs for involvement in other areas. The ideal situation would be an IASSIST membership for every professional employee of every data archive and data library. Those organizations with larger staffs typically have interests in the full range of IASSIST activities and could there- fore encourage participation of staff members in the work of every Action Group; smaller organizations could have staff members in a subset of the Action Groups.

The Classification, Data Archive Development, Data Organization and Manage- ment, and Process-Produced Action Groups already have large enough constituencies of active members to pursue their assigned tasks. The Data Archive Registry and Documentation Action Groups have made some progress, but need a larger base of members in order to make more substantial progress. At least in North America, Data Acquisition will temporarily be monitored by Data Archive Development and Process-Produced Data.

A report on IASSIST was presented at the Special Library Association annual meeting in New York City on June 6, 1977. IASSIST members are also actively in- volved in the organization and planning of an invitational conference on the cat- aloguing of machine-readable data files to be held in Washington in January 1978. Please keep me informed of additional opportunities for IASSIST participation in the meetings of other organizations and encourage your colleagues to send in their dues.

## ACTION GROUP REPORTS

### DATA ARCHIVE REGISTRY

Canada-  Lisa Lasko, Canadian Consortium for Social Research, Institute for
         Behavioral Research, York University, 4700 Keele Street, Downsview,
         Ontario M3J 1P3

Europe-  Joseph Bonmariage, Belgian Archives for the Social Sciences, University
         of Louvain, SH-2, 1348 Louvain-la-Neuve, Belgium

United States-  John Kolp, Regional Social Science Data Archive, University of
                Iowa, Iowa City, Iowa 52242


### Report of the Joint Canadian-United States Action Groups on
### Data Archive Registry Meetings in Toronto, May 11-12, 1977 (DAR AG)

Submitted by

Lisa Lasko
Institute for Behavioral Research
York University

Members Present

Canada
  Lisa Lasko, Institute for Behavioral Research, York University
  Sharon Henry, Data Clearing House for the Social Sciences
  Gerald Prodrick, University of Western Ontario
  Jana Prokop, University of Toronto

The Canadian Data Archive Registry Group made substantial progress when it
met for a working session at the Canadian IASSIST meetings, May 11 and 12, in
Toronto. The working sessions were attended by four Canadian IASSIST members.
No American members were present.

The first issue that was addressed by the group was whether or not the Data
Archive Registry Group still held a viable mandate. The presence of two new data
archive directories-- the upcoming 1977 September edition of Vivian Sessions'
Directory of Data Bases in the Social and Behavioral Sciences and the Unesco
sponsored Directory of Data Services now underway under the direction of Jean
Meyriat, Secretary General of the International Committee for Social Science In-
formation and Documentation (ICSSD) forced consideration of this issue. In ad-
dition, the Canadian Group had to consider that the Data Clearing House for the

Social Sciences intended to produce a hard copy <u>Directory of Social Science Data Centres</u> for Canada. The group agreed that the first two directories were less than satisfactory, but that the soon-to-be-produced directory of the Data Clearing House posed a real conflict. Sharon Henry, Executive Director of the Data Clearing House, reported that although she had to have the <u>Directory</u> published by the end of the current year, in fact, no work had yet been started on the publication and no questionnaire had yet been designed. It was suggested that a formal liaison between the IASSIST Data Archive Registry Group and the Data Clearing House be established to produce the directory. More specifically, the questionnaire utilized for surveying Canadian data centres would be jointly designed by both groups, and the data elements covered in the questionnaire would allow the results to be used by both groups for their various needs. One publication, the <u>Directory of Social Science Data Centres</u>, would be published cooperatively by the Data Clearing House for the Social Sciences and the IASSIST Data Archive Registry Group. However, it was clearly understood that when and if an IASSIST International Data Archive Registry did get underway, the relevant Canadian entries would be pulled from the data base for inclusion in the IASSIST registry. The Canadian directory, then, was to be viewed as a pilot project for the international IASSIST data archive registry. It was agreed that in order for this joint activity to work, a formal representative from the Data Clearing House for the Social Sciences should become a permanent member of the IASSIST Data Archive Registry Group.

Once this basic issue was resolved, the group immediately started to work on designing the questionnaire. Both the Session's questionnaire and the Meyriat questionnaire were consulted. By the end of the two days, most of the necessary elements had been agreed upon and Pierre Lacasse, coordinator of the Data Acquisition Group, made some valuable suggestions for elements concerning acquisition policies of archives. However, as work on the questionnaire did not get completed, it was decided that Lisa Lasko and Jana Prokop would meet in Toronto to complete it, after which time, a meeting would be held with Sharon Henry to discuss it further.

The group spent a substantial portion of its time compiling a list of broad subject headings to be used for describing categories of data in the questionnaire. During the conference, it was pointed out that the lack of a controlled vocabulary for descriptions of categories or holdings of data, was a major factor in the lack of good subject access to data archives. The group felt that its work in this area would be a significant and important contribution to useful descriptions of data archives. The group built upon the list of 105 subject headings compiled by David Gerhan and Loretta Walker, and also consulted with Vivian Sessions' 26 broad subject categories. In addition, unique Canadian subject terms were chosen. The final version differed substantially from the aforementioned lists, and, as it was still very rough, Lisa Lasko and Jana Prokop agreed to refine the list further at a later date in Toronto. Sue Dodd, Coordinator of the Classification Action Group, generously offered to assist in the compilation of subject terms. It was agreed that the refined list would be sent to Sue Dodd for comments and suggestions as soon as possible.

To summarize, then, the Data Archive Registry Group made significant headway in designing the questionnaire and developing a controlled vocabulary to be used for descriptions of data archive holdings. However, further work is still required to complete both the questionnaire and subject heading list. For this reason, Lisa Lasko and Jana Prokop both agreed to meet in Toronto in the next few weeks to take care of any work still outstanding.

## CLASSIFICATION

Canada- Mohan Sharma, Humanities & Social Science Library, University of Alberta, Rutherford North, Edmonton, Alberta

Europe- Ekkehard Mochmann, Zentralarchiv für Empirische Sozialforschung, Bachemer Strasse 40, 5 Köln 41, Federal Republic of Germany

United States- Sue Dodd, Data Library, Insitute for Research in Social Sciences, Manning Hall, University of North Carolina, Chapel Hill, North Carolina 27514

### Report of the Joint Canadian-United States Action Groups on Classification (C AG)

Submitted by

Sue A. Dodd, University of North Carolina

#### Members Present

Canada
  Martha Amschutz, Canadian Radio-Television Commission
  Krystyna W. Dynowski, University of Western Ontario
  Sue Gavrel, Public Archives of Canada, Ottawa
  Katrin Horowitz, Library of Consumer & Corporate Affairs
  Hans G. Schulte-Albert, University of Western Ontario

United States
  Sue A. Dodd, Ch, University of North Carolina
  Gertrude Lewis, Rutgers University
  Mimi Schade, Brookings Institution
  Philip Sidel, University of Pittsburgh

Agenda topics: Current problems and associated tasks within the mandate of the Classification Action Group include developing (1) examples and guidelines for bibliographic references for social science numerical data files, (2) a "cataloging-in-production" scheme for major producers of social science data files, (3) a more "universally based" classification scheme for social science data files; and, reviewing (4) the cataloging efforts to date and discussing any or all related problems and (5) existing printed thesauri in the social sciences in terms of their future applications to data files.

The primary focus of the C AG at the Canadian Working Conference centered around the first problem of how to cite properly a social science numerical data file in the published literature.

Problem: Currently, there are no standards or guidelines for compiling such bibliographic references and the quality and amount of information provided when citing social science data files varies greatly among individual scholars. In many cases, the information is not sufficient to allow for direct access, and consequently, an interested party has to spend a considerable amount of time determining additional information that could easily be provided by a full bibliographic reference.

Complicating this problem is the American National Standards Institute's forth-coming work entitled "American National Standards for Bibliographic References." This ANSI Standard will provide detailed information on the theory, principles, and definitions underlying the technique of preparing bibliographic references for both print and nonprint materials, including machine-readable data files (MRDF). The problem with this forthcoming document is that it does not include examples of social science numerical data files which make up the vast majority of MRDF; nor does it attempt to be compatible with the forthcoming second edi-tion of the Anglo American Cataloging Rules (AACR II), which also deals with bibliographic standards for MRDF.

Role of IASSIST C AG: The Classification Action Group, in its continuing role to encourage uniform standards for MRDF and to see that social science data files are represented in the best possible light will provide a written cri-tique on the forthcoming ANSI Standard, and at the same time provide more ser-viceable examples of how to cite social science numerical data files in the published literature. The C AG does not intend to create a new standard but rather to work within the framework of the ANSI Standard. The intent is to supplement it and recommend revisions where it is viewed as either necessary or helpful.

The critique will focus on the omission of social science numerical data files and the lack of compatibility with other standards and related terminology dealing with the same medium. Arguments to be presented in the paper are:

that outside the field of quantitative research, social science numeri-cal data files are probably the least understood category of MRDF and, consequently, are in the most need of attention and clarification;

that the terminology and definitions used to describe MRDF are confusing and often misleading and, therefore, a uniform standard for this medium is necessary;

that social science numerical data files make up the largest on-going collection of MRDF and, consequently, this body of information is in immediate need of bibliographic control;

that social science numerical data files are cited more frequently in scholarly journals than the two MRDF examples listed in the ANSI Stan-dard (i.e., computer programs and bibliographic data files), and there-fore, the need for relevant social science examples seems justified;

that this justification is magnified by the fact that the manner in which social science data files are cited in the literature plays an important role both in terms of access and in terms of secondary an-alysis; and,

that the two examples given (i.e., computer programs, and bibliographic data bases) are significantly different from the characteristics of social science numerical data files as to be less than helpful.

The first draft of this paper was reviewed by the participants of the Canadian Working Conference and will likewise be reviewed by those C AG participants who were unable to attend the conference. The final work will be reviewed by the C AG's Standards and Quality-Control Review Board. The resulting critique, along with an accompanying letter signed by the appropriate officers of IASSIST, will be sent to the American National Standards Institute, Inc. in New York and beyond that, it is hoped that the recommended revisions and resulting examples would be accepted by this body. It is also hoped that relevant examples could be sent to the editors and publishers of the various social science journals, who in turn can begin to implement conventions on the use, style and content of bibliographic references.

Tasks of C AG participants: (1) read and review the draft of the critique of the ANSI Standard, providing any comments, suggestions, or additions, etc.; (2) read an abbreviated version of the standard prepared by Ellis Mount and which appeared in the Journal of the American Society for Information Science; (3) prepare a list of the necessary and descriptive bibliographic elements (e.g., title, author, edition, imprint, etc.) for social science data and at the same time indicate those that would be essential, recommended, or optional; and, (4) utilizing the list of rated bibliographic elements, begin to compose bibliographic references of the respective data files.

Results of this task at the Canadian Conference: The first draft of the critique was reviewed and comments accepted. The beginnings of a list of bibliographic elements was compiled. Each element was grouped and rated (i.e., essential, recommended or optional) according to its importance in identifying a data file. Finally, some eleven examples were compiled by the participants, including the samples listed below:
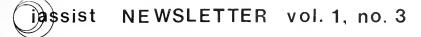
Title- Mexico's naturalized citizens, 1828-1931 [Machine-readable data file].
first: Harold Sims; Susan Sanderson; Philip Sidel, principal investigators.
Pittsburgh, PA: University of Pittsburgh, 1975-76. 1 data file (8066 logical records).

Author- Shanas, Ethel. The health of older people [Machine-readable data file]:
first: A social survey: public attitudes on older people. NORC rev. ed.
Chicago: National Opinion Research Center, 1957 [producer and distributor]. 1 data file (2567 logical records) and accompanying codebook (166 p.).

Swidzinski, Susan. Syllabication [Machine-readable data file]: a drill and practice lesson. Bloomington, MN: Control Data Corporation, 1976. On-line program lesson available only via the Plato System.

Henry, Neil. MAXCLS.BAS [Machine-readable data file]: a program for maximum likelihood estimation of parameters of unrestricted latent class models. Lafayette, IN: Gary Income Maintenance Experiment, 1974; Pittsburgh, PA: Social Science Computer Research Institute [distributor]. 1 program file (95 statements, BASIC) and accompanying manual (53 p.).

The work at the Canadian Conference on this task was an important first attempt towards achieving the goal of providing relevant and acceptable examples of bibliographic references of social sciences data files. However, these examples should be viewed as only first attempts and not as the final recommended examples of bibliographic references; and, any and all comments are welcome.

DATA ARCHIVE DEVELOPMENT

Canada- Laine Ruus, Data Library, Computing Centre, University of British Columbia, 2075 Wesbrook Place, Vancouver, British Columbia   V6T 1W5

Europe- not activated

United States- Alice Robbin, Data and Program Library Service, 4452 Social Science Building, University of Wisconsin-Madison  Wisconsin  53706


Report of the Joint Canadian-United States Action Groups on
Data Archive Development Meetings in Toronto, May 11-12, 1977 (DAD AG)

Submitted by

Laine Ruus, University of British Columbia

Members Present

Canada
  Laine Ruus, Ch, University of British Columbia
  Marilyn Berry, University of Victoria
  Peter Clinton, Memorial University of Newfoundland
  Judy Demaine, University of Guelph
  Ed Hanis, University of Western Ontario
  Alan Kirby, Queen's University
  Elaine Kozak, Data Clearing House for the Social Sciences

United States
  John Heddesheimer, National Archives and Records Service, Washington, D.C.
  Alice Robbin, University of Wisconsin-Madison
  Judith Rowe, Princeton University
  Ed Vickery, Research Triangle Institute, Chapel Hill, NC

Objectives of the meeting: (a) to revise and redefine the outline of the Guide to providing social science data services; (b) to elaborate and expand on the contents of the separate sub-sections of the Guide; (c) to finalize arrangements for contributions from the other action groups to relevant sub-sections of the Guide, and set deadlines for abstracts and full-text contributions; (d) to allo- cate responsibility for contributing remaining sections and sub-sections of the Guide to persons or groups with relevant expertise.

The first priority of the action group was to write abstracts of those sections of the Guide to be contributed by other action groups, so as to afford them a chance to consider more fully, during their meetings, their contributions and react to them. On the basis of these reactions, some sub-sections of the out- line will be redefined, and some responsibilities reallocated.

Specific responsibilities for remaining sub-sections of the Guide were allocated
to such persons of expertise in relevant areas as could be identified by the mem-
bers of the action group, both within and outside the action group itself. Those
who could not be contacted during the course of the conference, or who had not
previously been contacted, will be solicited for their collaboration immediately.

A deadline for contribution of abstracts of all sections was set for June 30, 1977,
and deadline for full-text for December 31, 1977.

In addition, a system of readers, for the several sections of the Guide, was es-
tablished to provide a pre-editorial review of input, and to ensure that foci of
the sections include minimal as well as maximal levels of service.

It was further decided that preliminary copies of the glossary should be dis-
tributed to all contributors and readers for comment and ammendment, to ensure
that all vocabulary necessary is included, and that definitions in the glossary
agree with usage by the contributors.

Based on reaction from members of the Data Organization and Management Action
Group, it was found to be necessary to further refine and amplify the defini-
tion of the target audience, especially in terms of levels of expertise.  A
fuller redefinition will be distributed to all collaborators.

The final order of business was a decision on a topic for the DAD-AG symposium
at the February 1978 meetings.  Various topics were considered, and the final
choice was 'networking', as a topic of some immediacy for data archives/librar-
ies.  Consideration will be given to a definition of 'networking', and its im-
plications for all machine-readable data service units.

## PROCESS-PRODUCED DATA

Canada-  John DeVries, Social Science Data Archives, Department of Sociology,
          Carleton University, Ottawa, Ontario  K1S 5B6

Europe-  Paul Müller, Institute for Applied Social Research, University of
          Cologne, Greinstrasse 2, 5000-Koln 41, Federal Republic of Germany

United States-  Donald Harrison, National Archives (NNR), Washington, D.C.
          20408

Report of the Joint Canadian-United States Action Groups on
Process-Produced Data Workshop in Toronto, May 11-12, 1977 (PPD AG)

Submitted by

John DeVries, Carleton University

Members Present

Canada
  John DeVries, Ch, Carleton University
  Bill Bradley, Health and Welfare Canada, Ottawa
  Hy Burshtyn, Carleton University
  Tony Falsetto, Public Archives of Canada, Ottawa
  Richard Guttormson, Ministry of State for Science and Technology, Ottawa
  Pierre Lacasse, Centre de recherche en aménagement régional, Sherbrooke, Québec

United States
  Harriet Dhanak, Michigan State University
  Shirley Gilbert, Princeton University
  Elizabeth Powell, LEAA/NCJISS, Department of Justice, Washington, D.C.
  Ed Vickery, Research Triangle Institute, Chapel Hill, NC

We continued our discussions on the Directory of Catalogues which list Machine
Readable Data, which we had determined to have the highest priority for our AG.
Although we did not meet our self-imposed deadline (which expected us to produce
a first draft of this proposed document at this meeting), some progress has been
made.  We have obtained copies of several existing catalogues and propose to con-
tinue this collection phase.

We developed a listing of minimally required elements of information, which we
hope each entry in a Catalogue of Data Files would provide.  The listing which
follows was based on an examination of existing catalogues.  The essential items,
we feel are:

    a)  File name, catalogue number or reference number (if any);

    b)  A short description of the data.  This description should especially indi-
cate the time-span and the geographical area covered by the data;

    c)  Medium on which the data are held (e.g. magnetic tape, cards);

    d)  Size of file (number of records);

    e)  References to available documentation and other publications pertaining to
the file (e.g. a listing of reference numbers for working papers);

    f)  The degree of accessibility (we propose:  unconditional, conditional, never);

    g)  Cost of acquiring the file (if the file is accessible);

    h)  Department or agency where the raw data originates;

    i)  The position of the person to contact for further information.

On the last-mentioned item, we felt that the name of the person, and the telephone number, were additional items of importance. Given the high degree of turnover of personnel, we felt that the position would generally be more enduring information than the name of the incumbent at a given point in time.

We envision the eventual document--still planned to appear in February 1978--to be an annotated bibliography of Catalogues, where the annotations are based on our proposed set of minimal information requirements. The bibliography would also state these requirements, and would, finally, contain a listing of governmental departments and agencies which, to our knowledge, have not (yet) produced a cata- logue of their machine-readable data files.

We are proposing to circulate our initial findings to the members of our AG before August 31, 1977. At this point each AG member will be asked to produce the anno- tations for a specified set of catalogues. We expect to review the eventual set of annotations at the projected IASSIST meeting for February 1978.

With regards to our second priority: the Inventory of Existing Guidelines in Use by the Originating Institution or Archival Institution, we decided that little could be done at this stage, until we had an opportunity to examine the Catalogues of Data Bases. In addition, Tony Falsetto will mail out (again) a package con- taining various documents related to the documentation and preservation of MRDF by the Machine-Readable Archives Division of the Public Archives of Canada.

We discussed the issue of communications between national governments and the com- munity of researchers using process-produced data. We agreed that there is a great need for improvement and propose the following strategy:

(1) Comments, suggestions, queries and complaints about process-produced data would be sent to an IASSIST member who is at the same institution as the person requiring the information (or in a nearby institution). In short, we would pro- pose that IASSIST endorse a network of "Official Representatives".

(2) Such questions, etc. would then be passed on to the appropriate AG coordin- ator (i.e., the U.S. Coordinator if one were dealing with U.S. PPD, the Canadian coordinator in the case of Canadian PPD).

(3) The AG coordinator would convey the question to the appropriate government agency and would be responsible for channeling the response back to the user (or, if the information were of wider relevance, to submit an item to the IASSIST News- letter).

The AG also discussed ways in which other organizations working on related problems could be contacted to prevent any duplication of effort regarding process-produced data. We agreed that the two coordinators would get in touch with the chairman of such groups as the Association of Public Data Users, the Committee on Population Statistics of the Population Association of America, and so on.

DATA ORGANIZATION AND MANAGEMENT

Canada- Greg Morrison, Social Science Data Archive, Department of Sociology, Carleton University, Ottawa, Ontario K1S 5E6

Europe- Eric Tannenbaum, Social Science Research Council Survey Archive, University of Essex, Wivenhoe Park, P.S. Box 23, Colchester, Essex, England CO4 3S0

United States- William Gammell, Social Science Data Center, University of Connecticut, Storrs, Connecticut 06268

Report of the Joint Canadian-United States Action Groups on
Data Organization and Management Meetings in Toronto, May 11-12, 1977 (DOM AG)

Submitted by

Greg Morrison, Carleton University

Members Present

Canada
  Greg Morrison, Ch, Carleton University
  Clement K. M. Chan, McMaster University
  Janet Chan, University of Toronto
  Rachel Des Rosiers, Data Clearing House for the Social Sciences
  Robert Logan, University of Guelph
  Paula Mitchell, University of Western Ontario
  Richard L. Schnaar, Public Archives of Canada, Ottawa
  Terry Stewart, University of Waterloo
  Richard Wolfe, Ontario Institute for Studies in Education

United States
  Bill Gammell, Ch, University of Connecticut
  Barbara Aldrich, University of Wisconsin
  Gary M. Grandon, University of Connecticut
  Pnina Grindberg, Columbia University
  Sheldon Laube, C. M. Leinwond Associates
  Barbara Noble, University of Illinois
  Sharon Poss, Duke University
  Richard Roistacher, University of Illinois
  Phil Sidel, University of Pittsburgh

During the Toronto sessions, members of the Data Organization and Management Group engaged in a reassessment of how various sorts of activities might best be carried out. It was felt that some kinds of projects will have to be done essentially by individuals or small sub-groups, in accordance with their particular areas of interest and expertise. Draft documents, papers, proposals and the like will then be sent to the co-ordinators of the Group, who will circulate them to Group members for reactions and comments. After a review of possibilities, various activities were taken on by individual members.

(a) The most important activity of the Group during the next year will be the writing of sections of the Guide to Providing Social Science Data Services. In consultation with a member of the Data Archive Development Action Group, eleven people agreed to draft material, and another two volunteered as sub-editors.

(b) The development of a list of recommendations, addressed to potential researchers and data collectors, concerning study design as it relates to data management (the "Do's and Don'ts List"), first considered during the Group's sessions in Florida, has now been assigned to Rachel Des Rosiers and Terry Stewart.

(c) The proposal for an Inventory of Software relating to data management, also initiated in Florida, has been rather drastically revised, in view of the existence of other groups outside IASSIST who are reviewing social science software, the need for clearer definition of what kinds of functions and programs ought to be included, and the formidable amounts of time, energy and other resources required for a comprehensive job. It seemed wiser to adopt an incremental approach and start with two clearly delineated projects: a survey of the programs used at one institution (Columbia University), to be undertaken by Pnina Grinberg as a pilot study; and an investigation of methods and available programs for cleaning multipunched data, to be done by Bill Gammell, Gary Grandon and Greg Morrison. Other such narrowly-focused efforts may be undertaken in the future.

(d) The Group felt it would be useful to gather information about the areas of technical expertise of the members of IASSIST, and to organize and publish the results. The aim would be to facilitate the use of the IASSIST membership as resource people or consultants for each other on technical problems. As a by-product, a profile of the members could be produced. The best procedure for obtaining the required information is under discussion. Gary Grandon and Sheldon Laube will be carrying out this activity.

(e) Another area of interest is the holding of workshops. Sheldon Laube will consider possibilities for organizing workshops at future IASSIST meetings, while Bill Gammell will investigate possibilities for regional workshops.

(f) Consideration of the Inventory brought home to the Group the importance of establishing communication with other organizations working on related problems. The co-ordinators, Bill Gammell and Greg Morrison, will conduct these external relations.

In addition to the above projects, for which individuals have taken on responsibility, two other issues were discussed. The first was the problem of information dissemination and exchange, crucial in the area of data organization and management, where relevant literature appears in a very wide range of publication from many different disciplines. We feel that this problem is particularly amenable to action by all IASSIST members, whether or not they are affiliated with the AG. We urge IASSIST members to send the co-ordinators of the Data Organization and Management Action Group and/or the Editor of the IASSIST Newsletter relevant bibliographical references to any material which relates to data handling, broadly defined. These references could be to published articles or unpublished papers, to books, technical reports, computer programs, items in newsletters, and so on. The reference should include a standard bibliographical citation, information on where to obtain the item if not found in readily available sources, and preferably an annotation or abstract. This information will then appear in the Book Notices section of the Newsletter. Also welcome would be short technical notes, resumes of procedures, documents, etc.; these ought to be sent to the co-ordinators for scrutiny and possible publication in the Newsletter.

The second general issue discussed was that of IASSIST endorsements. While the Group could see the value of IASSIST lending its support to projects in progress, it was not clear whether or not this should come as an official action and if so what process should be followed in determining what to endorse. A distinction was drawn between endorsements of good practices and endorsements of particular products; the latter possibility raises questions which might be considered by other Action Groups or by the Steering Committee. Short, however, of formal endorsements, members can provide feedback for the originators of projects, and engage in lobbying activities on an individual basis on its behalf.

## DATA ACQUISITION

Canada- Pierre Lacasse, Centre de recherches en aménagement régional, Université de Sherbrooke, Sherbrooke, Québec

Europe- Marcia Taylor, Social Science Research Council Survey Archive, University of Essex, Wivenhoe Park, P.O. Box 23, Colchester, Essex, England CO4 3S0

United States-

[No report submitted.]

## DATA DOCUMENTATION

Canada- Dave L. Salley, Management and Central Services Group, Standards Division, Statistics Canada, Tunney's Pasture, Ottawa, Ontario K1A OT6

Europe- Cees Middendorp, Steinmetzarchief, Kleine-Gartmanplantsoen 10, Amsterdam-C., Netherlands

United States- John Grasso, Office of Research and Development, Center for Appalachian Studies and Development, West Virginia University, Morgantown, West Virginia 26506

[No report submitted.]

## IASSIST ESTABLISHES PUBLICATION POLICY

### Publication Policies for IASSIST

A. Copyright (including royalties) and control of publications produced by
   IASSIST, via action groups and other professional mechanisms, rest with
   IASSIST, subject to terms that might be negotiated with cooperating organ-
   izations (e.g. ALA) and publishing houses.

B. For each publication approved by IASSIST, the principal author(s) and/or
   editor(s), depending on the circumstances will be identified, and an agree-
   ment will be struck in which proprietary interests are conveyed to IASSIST
   in return for:

   1. Professional recognition to the principal worker(s) in the form of au-
      thorship or editorship of the work as may be appropriate.

   2. A share of royalties that might be earned beyond some threshold level.

   3. Acknowledgement of all members of the working party which participated
      in the preparation of the work.

C. IASSIST will undertake the development of a publication strategy for the
   publication of its professional outputs.  The publication strategy will in-
   clude:

   1. Timing and priorities

   2. Identification, coordination, and integration of the outputs of its
      various working parties.

   3. Negotiations with other organizations for joint publications.

   4. Negotiations with publishers.

   5. Plans for successive stages of publications through the IASSIST News-
      letter, conference proceedings, articles in relevant journals, manuals,
      and books.

D. Publication Policies will be administered by an Editorial Review Board
   composed of five members (to be appointed by the Steering Committee).

**PAPERS PRESENTED AT THE SECOND IASSIST NORTH AMERICAN MEETING, MAY 11 – 12, 1977**

[Editor's Note: The following are abstracts of papers presented at the Second IASSIST North American meeting. The full proceedings of the Conference will be available in the near future. All participants of the Conference will receive a full copy of the Proceedings; IASSIST members will receive a copy at a reduced rate; all other individuals should contact Sharon Henry of the Data Clearing House for the cost to non-IASSIST members.]

IASSIST - Progress Towards Solving Problems in Data Archiving

Carolyn Geda
Inter-university Consortium for Political and Social Research

The paper reviews the establishment and development of IASSIST including the rationale for such an Association. Problems encountered during the formation of the International Organization are presented with interim proposals for dealing with them. The governing body, regional structure, and functions of the action groups are discussed. A brief review of Regional activities is given and potential directions for the Association are projected.

Canadian Secretariat Report

Rachel Des Rosiers
Data Clearing House for the Social Sciences

The report summarizes the membership campaign for IASSIST in Canada. It outlines the results as well as expected plans to increase the membership in Canada, e.g., by contacting related associations in the field and publicising IASSIST information in professional journals and at conferences. It also discusses its relationship with the American secretariat, especially in regard to Action Groups' projects and plans. Finally, it invites comments and suggestions from the Canadians present concerning its intended role and functions.

IASSIST Secretariat Report: United States

Judith Rowe
Princeton University

The report of the United States Secretariat covers the current status of membership enrollment as well as proposed membership recruitment activity. Specifically, it addresses the goal of recruiting for Action Group activity all of the professional staff of each data archive and data library. It includes a general summary of both secretariat and Action Group activity, as well as a report on the first North American Action Group Conference in Cocoa Beach. Other topics include an initial report on the plans for next year's conference and some activities and programs planned by other organizations which would be of interest to IASSIST members. Among the latter are the annual conferences of Special Libraries Association (SLA), Association of Public Data Users (APDU) and the American Association for Information Science (ASIS), as well as the Inter-university Consortium for Political and Social Research (ICPSR) workshop for data librarians.

Data Archive Registry-Survey of Past Effort,
Suggestions for the Future

Lisa Lasko
Canadian Consortium for Social Research

The mandate of the IASSIST Action Group for Data Archive Registry is described. The major portion of the paper is devoted to a brief review, description, and evaluation of some of the more important existing directories. These are: Social Science Data Archives in the United States, published by the Council of Social Science Data Archives in 1967; A Directory of Information Resources in the United States: Social Sciences, 2nd edition, published by the Library of Congress in 1973; the second edition of the Encyclopedia of Information Systems and Services, edited by Anthony Kruzas and published in 1974; and the Directory of Data Bases in the Social and Behavioral Sciences, edited by Vivian Sessions and published in 1974.

Recent developments in data archive registries, such as the UNESCO sponsored Directory of Data Services and the Directory of Data Centres, to be published by the Data Clearing House for the Social Science in Canada are also discussed. Essential elements and user requirements of social science data archive directories are dealt with in reference to these directories.

Finally, a summary of the problems facing the Data Archive Registry Group, both abroad and in Canada, is given, along with several proposed options include a discussion of the viability of a Data Archive Registry Group in Canada; unique contributions to be made in this area, one of which might be a compilation of a list of subject headings for social science data archives; and, the production of a directory of data archive personnel.

Problems in Handling Process-Produced Data

John DeVries
Social Science Data Archive
Carleton University

The paper is an attempt to outline the problems one encounters in the handling of process-produced data, starting with their definition through the stages of obtaining information about them and acquiring them, to the phases of transforming and analysing them. In addition to enumerating and discussing the often unique problems with which the user of process-produced data is faced, I will try to mention various projects which have already begun, and indicate future plans by the IASSIST Action Groups on process-produced data.

Definitions. In the last few years, various definitions have been suggested. Although there is considerable overlap between them, the remaining differences make it important that we come to some agreement about the proper "domain" for an Action Group on process-produced data, and thus to a clear definition which unambiguously specifies this domain. I will discuss the definitions of which I am aware, and present some arguments in support of the definition which was developed at the Cocoa Beach conference.

Information and acquisition. Generally, we will find that process-produced data have been generated by governments and other public organizations. Problems in this phase are thus to a large degree subsumed under the general heading: "The relationship between the research community and the government." Specific issues I will discuss include: privacy versus the right to information, i.e., the government as a research resource; and, obtaining information from governmental agencies.

Documentation. In many non-trivial ways, process-produced data differ from surveys. Some of the unique aspects regarding documentation are: definitions of the universe to which the data pertain, description and estimates of the errors of coverage and content, and requirements for a "codebook".

Data transformation. Once a data set has been acquired by a researcher or an archive, there are various difficulties one faces in transforming and analyzing the data. Many data sets are "ragged" rather than rectangular, thus requiring either intricate transformations or specialized statistical packages. Aggregated data require a high level of statistical sophistication on the part of the analyst. Finally, one will frequently wish to link data from various data sets. In doing so, one often runs into problems of comparability: stimuli may not be standardized across data sets; for identical or comparable stimuli, response categories are frequently not identical across data sets; different data sets may not relate to exactly identical universes; finally, where spatial and/or temporal delimiters are involved, they are often not identical across data sets.

Data Acquisition for Archives

Pierre Lacasse
University of Sherbrooke

The purpose of this paper is to raise a number of issues about data acquisition in order to narrow down the subject to the most important avenues for further discussion. With so diversified stock of social data being created, we need to ask "who will keep what, and on what basis"? To clarify this question we will particularly discuss three major inputs to the subject: (1) The collectors of data: Who are they? What types of data are they collecting? What data are we interested in? (2) The archives: What are they? What are their purposes in holding data files; and for what types of secondary users? What mandate do they have and what is their scope of interest? (3) The users for whom data is held: Who are they? Where do they stand? On what issues? What are their needs? How will they use the data held by the archives?

After discussing these questions, we will take a look at what has been done in the data acquisition action group. In particular we will discuss the questionnaire prepared by the European section of IASSIST on "Archives data acquisition policies and problems."

Consider the group's mandate, which requires proposing ways of linking collectors and archives, with the stress on defining the group objectives and conditions which must be met in order to close the gap between collectors and archives. Finally, we will try to suggest some possible courses of action in the short and long term.

Organization of Data Archives

Laine Ruus
University of British Columbia

The mandate of the IASSIST Data Archive Development Action Group is two-fold: the creation of a "procedures manual consolidating current archival organizational, administrative, and personnel structures, procedures, and policies...to aid developing archives", and the organization of training workshops and seminars to aid in personnel training and professional development. The efforts of the Data Archive Development Action Group have been entirely concentrated on the development of A guide to providing social science data services. While response to past workshops has indicated the need for this guide, plans are for the moment in abeyance. Problems in data archive development occur at three levels. Intra-archival problems are primarily administrative: planning and policy making, staff, acquisition, data management, technical and user services. Intra-institutional problems, impinging on the former, occur in the areas of planning, inter-departmental cooperation, support services, and funding. Problems at the inter-archival level result primarily from lack of adequate standards and conventions and from underdeveloped formal channels of information dissemination.

Standards for Data Documentation

Dave Salley
Statistics Canada

In recent years, a number of initiatives related to the concept of data documentation have been undertaken within the Canadian Federal Government. The motives for these initiatives have centered around the need to inform users on data availability, the need to monitor duplication of data collection, the need to control questionnaire content and the need to assess the costs incurred in the collection, compilation and dissemination of federal information. While these motives are not exhaustive, they do represent a significant class of problems for consideration by the IASSIST Action Group.

The shortcomings of most efforts to date have centered on the particular or special purpose nature of the systems. In many cases, a particular subject area or class of users restricted the approach severely. In short, little consideration has been given to the development of standard components of a Data Documentation system with a view to meeting general requirements. The reason for this problem may well be that information managers in the Federal Government have not yet developed a real appreciation of the need for standard documentation procedures, both as a tool for managers and planners of information activities and to assist "end data users". Thus, most initiatives have been quite ad hoc and narrow in scope. However, one should not underestimate the useful technical approaches and methodologies developed to date.

As a departure point for the Action Group, it is suggested that Terms of Reference be developed which contain a strong "Standards" ingredient. Furthermore, an intensive needs and benefits study should be undertaken with respect to the whole data documentation issue. Certainly, the group should bring together current work in a systematic fashion, but the development of standard approaches in the scope, technique, uses and purposes of data documentation systems will have far more useful results in the long run.

Cataloguing and Classification of Machine-Readable Data Files;
A Preliminary Report on the US IASSIST Classification Cataloguing Project

Sue Dodd
University of North Carolina

The primary emphasis of the US Classification Action Group of IASSIST has been on establishing standards and on the study of library information systems as they may apply to social science data files. Some of the recent developments within the library system which the Classification Group is examining include: 1) the development of rules and guidelines for cataloguing machine-readable data files (MRDF); 2) the development toward the acceptance of the MARC (Machine-Readable Catalog) record format as a universal standard for the automated bibliographic record; 3) the development of networks and on-line information systems which allow for multiple input and immediate retrieval of information; 4) the development of thesauri and controlled vocabularies for social science terms; and, 5) the development towards future considerations of a national union list of available MRDF and their location.

Given the current work and interest in cataloguing MRDF, the first task of the IASSIST Classification Group was to participate in an organized project designed to test the feasibility of cataloguing social science data files according to the American Library Association (ALA)'s designated Subcommittee to recommend rules for cataloguing MRDF.

To facilitate the test, which was conducted by mail, a "Working Manual for Cataloguing Machine-Readable Data Files" was compiled based on an interpretation of the ALA Subcommittee's recommendations. The task required that the participants select six data files, either numerical, text or program files, and proceed to catalogue these data using the information and guidance in the manual; apply subject descriptors on the content of the files; and complete an evaluation form. The outline of the project, the actual test, and the preliminary results will be discussed in this paper.

<div align="center">

Data Organization and Management
Applications in Data Archiving

Greg Morrison
Social Science Data Archive
Carleton University

</div>

The formal mandate of the Data Organization and Management Action Group of IASSIST is used as the starting point for consideration of various kinds of problems with which the Group might concern itself -- problems which most data archivists will confront sooner or later. These include: the transfer of datasets; the cleaning, editing, transforming, merging and sub-setting of data; the organization of data files for the above activities and for statistical analysis; the technical aspects of the documentation of individual datasets by codebooks; and, the technical aspects of the documentation of collections of datasets by catalogues or inventories.

While an obvious organizing focus of the Group is computer software, it will try to define appropriate procedures in these areas, and to encourage the exchange of information about programs and procedures. The activities of the Group to date are reviewed, and some suggestions are made for the future.

**BOOK NOTICES / kathleen m. heim**

Lucci, York; Rokkan, Stein; and Meyerhoff, Eric. <u>A Library Center of Survey Research Data: A Report of an Inquiry and a Proposal</u>. New York: Columbia University School of Library Service. June 1957. [Available in microform from the: Columbia University, School of Library Service, New York, New York 10027]

It has been twenty years since the Lucci and Rokkan report first appeared. In many ways the data archive movement may be said to have begun with its publication. Many ideas presented therein have been built upon by later archivists and scholars. It is particularly interesting that while the Columbia Library School received the grant for the report, libraries have, until quite recently, been reticent to become involved in data archive activities. [See review of Howard Dalby White's Ph.D. dissertation below.]

### Background of the Report

Rapid growth of opinion polling and survey research following World War II made cross-national research a real possibility for scholars through the use of secondary analysis of comparable information collected in different countries. The difficulties of locating and obtaining even domestic researches, however, proved discouraging. The need for a library of survey research data which would assemble the more important survey data and make them readily accessible to scholars for secondary analysis prompted Professor S. M. Lipset then at Columbia University to propose that an investigation be made of the need for and the problems of establishing an international library center of survey research materials. In response to this proposal the Behavioral Sciences Division of the Ford Foundation awarded a grant in the spring of 1955 to the Columbia University School of Library Service to undertake such an investigation. York Lucci was given responsibility for the investigation in the U.S. and Stein Rokkan for Europe. In addition, Eric Meyerhoff served as a consultant on the more technical archival considerations.

The field investigators were charged with ascertaining and evaluating four areas: 1) Potential utilization of a center for research materials; 2) Availability of research material, its cost and difficulties in acquiring; 3) Adequacy of available data in terms of their potential for secondary exploitation; 4) Research and administrative requirements involved in such a library center.

The results of the Lucci U.S. investigation and Rokkan European investigation appeared as separate reports since the situations in the U.S. and Europe differed greatly with respect to the availability of survey materials and the extent of interest among scholars.

## Part I.  The Situation in the U.S. and Recommended Action

### Background and Purposes

Lucci discusses two chief reasons for a centralized archive:  lack of funds for the collection of primary data and inadequate use of data already collected. He then outlines past efforts to collect and disseminate results of polls and surveys such as the 1938 Public Opinion Quarterly's AIPO poll results, various periodical compilations published for short intervals, and the Roper depository at Williams College.  He concludes that published releases do not permit systematic analyses and that any archives that exist have been too limited and have failed to develop archival practices permitting usage by a wide community of scholars.

The "Loss of Data" receives much attention in the Lucci report.  A salient example is that of the Canadian Institute of Public Opinion where all but eight of 212 polls between 1941-1951 had been destroyed.  With other examples of data loss, such as the experiences of the Italian DOXA organization, the Swedish Gallup Institute, and the U.S. Office of War Information, Lucci builds a strong case for the need to undertake a strong preservation effort.  Parallel developments in other disciplines of basic tools for research (cyclotrons for scientists, libraries for historians) are compared with the lack of basic preservation of the raw material for social science research.  The Human Relations Area Files is discussed as the best known effort in the social sciences to assemble information from a variety of sources and make it available to a wide community of scholars.

One hundred and thirty seven social science researchers in the U.S. were contacted for ideas and reactions concerning a central data library for the social sciences.  Their ideas are consolidated and set forth by Lucci in the following sections.

### Sources of Data

The massive accumulation of data gathered by universities, the government, independent agencies, and foundations cannot be readily documented, but Lucci estimates that five to ten million interviews have been conducted annually in the post-wars years.  One market research organization alone reported 25 million cards in storage.  Though much of the data may be insignificant, if even five percent offer the possibility of scientifically useful secondary analysis, a quarter to a half-million cards a year might be stored.  Urgently needed is some means of locating that fraction of theoretically significant material, assembling it, making its contents available to researchers, and encouraging its exploitation.

Queries to survey organizations and academic institutions produced an overwhelmingly positive reaction about sharing data.  Lucci details at length responses from prominent centers, polling agencies, governmental agencies, and commercial research organizations and includes a section on "Conditions Imposed by Contributors."  The main objections he found to a proposed center were feelings that one cannot work successfully with someone else's data unless there is

close liaison with the person or agency initially responsible for collecting the data. Lucci counters this by advocating complete documentation and noting that "if survey research is to make any claim to science it can hardly refuse, indeed it should welcome, having a set of facts subjected to close scrutiny and the possibility of alternate interpretation.

## Utilization of an Archive of Survey Data

Uses of survey data for secondary research are discussed and six major points elaborated upon: 1) as the only primary source of knowledge about certain facts or patterns of behavior; 2) as a means of testing hypotheses; 3) the cumulation of cases; 4) to prepare for new primary research; 5) to assist the work of the historian; and, 6) to train students.

Objections to the proposal for establishing a center of survey research materials related to the extent to which utilization could be anticipated. Respondents to Lucci's questions felt that technical familiarity with methods of quantitative research was not widespread and that only large universities such as Michigan, Columbia, and Chicago could make use of such facilities. Lucci feels that availability will increase the skills of the entire social science community, and then reports comments of respondents, summarizing at length areas where respondents felt the center could be especially useful.

Further justification for the establishment of a center is given by a bibliography (pp. 134-38) appended to the report of published works based on secondary analysis, which represent different ways in which secondary analysis techniques can be used; expansion upon the use of secondary analysis by students and commercial organizations; and, the need by international organizations such as UNESCO for access to such a data center.

## Conclusion and Recommendations

Lucci summarizes by noting, "On the basis of the present inquiry, it can be stated flatly that the data are available...and it seems fairly clear that the possibility of their use is present and can be developed further." Such a center would maximize the use of available social science talent (especially at universities with minimal facilities); maximize the use of research funds; promote the comparability of data collection; facilitate cross-national research; preserve data in a systematic manner; and, spearhead new studies where lacks are noted.

A "Specific Proposal" is outlined with functions of the center noted at length: systematic collection and preservation of data; an index to data stored; information activities about the data to be disseminated to the research community; promotion of the center's materials; provision of duplicate sets of data for researchers with their own facilities; provision of tabulations where appropriate; maintenance of records of secondary research activities and publications; and, promotion of training and standardization.

Each of these functions is discussed in great detail along with possible locations for the center, staff considerations, equipment and storage needs, operating costs and sources of funding.

### Part II: A Review of the Situation in Western Europe and
### Some Recommendations for Possible Action

Rokkan sketches rough estimates of the data accumulation problem in Western Europe. He estimates that in the four major nations: France, Italy, the United Kingdom and Germany, half a million interviews have been generated in each nation since 1951. Among the smaller countries: Austria, Belgium, Denmark, Finland, the Netherlands, and Switzerland, the yearly production of interviews is on the average 75-100,000. Thus, the total interviews for Western Europe are between 3 and 4 million annually.

The problems of preserving, storing and choosing relevant data from this quantity is seen by Rokkan as directly linked to problems of social science utilization. Most polling agencies limit analysis to the presentation of overall response distribution, never analyzing the data as thoroughly as their methodological quality and their theoretical relevance seem to justify.

Surprisingly little has been done by Western European social scientists to avail themselves of these analysis opportunities. The lag in utilization reflects the conservative reluctance of great numbers of social scientists to accept the new techniques as scholarly and scientific, the uneven distribution of statistical analysis skills among academic social scientists, and the difficulty of access to the primary materials in the polling agencies.

During 1955-56 Rokkan discussed problems of storage and data utilization with directors and officers of survey research organizations and with university teachers and research workers in twelve countries of Western Europe, as well as at regional meetings such as WAPOR, ESOMAR, the Congress of the International Political Association, and the International Sociological Association, in order to elicit reactions to the possibility of a solution through the development of an international archive of survey materials.

The possibility of an international archive was presented to survey organizations by a questionnaire devised by Rokkan. The response was unanimous in emphasizing the need for systematic encouragement of secondary analysis, although opinions varied about the steps to be taken to facilitate secondary analysis of data already assembled.

The responses were capsulized by ESOMAR President, Leif Holback-Hanssen: Western European materials should be stored and classified in a separate facility located within the Western European region; development of the archive should be planned in close association with representatives of the survey profession; and, the archive should be built up gradually through a series of pilot analyses.

Rokkan's queries to academic social scientists are discussed in the context of the status of survey research in Western Europe: i.e., there was little being conducted. Most social scientists tended to favor independent gathering of data in order to build gradually to the need for a regional or international repository. Rokkan then discusses the format of pilot analyses and possible topics of these analyses.

Rokkan concludes: there is a need for fuller utilization of the bodies of information assembled by survey organizations through systematic analyses from the primary record; a need for the organization of a program of comparative secondary analyses of survey materials in selected substantive areas and, concurrently with them, the development of an archive of selected materials; and, a need for "groundwork" analyses and an archive within the Western European region. Cost estimates are given.

## Appendices

As noted above, a selected bibliography of works done using secondary analysis techniques is included. Another appendix details methods for indexing and cataloguing archival materials.

## Reviewer's Comments

The Lucci-Rokkan report is a fundamental document for the data archivist. The well formulated report gives the background of the data archive movement and predicts many of the developments we are now experiencing. This report is essential reading for all involved in data archives, because of the perspective it provides on the U.S. and Western European research communities and because of its visions of the future.


White, Howard Dalby. "Social Science Data Sets: A Study for Librarians." Ph.D. dissertation, University of California, Berkeley, 1974.

As noted in the review of the Lucci-Rokkan report above, libraries have not generally accepted responsibility for acquiring data in machine-readable form although a number of scholars and librarians have argued that machine-readable data is a logical extension of a library's collection. White explores the relationship between libraries and archives, giving a historical overview of arguments that advocate the placement of data archives within libraries. He compares library and archival functions and although he explains that many tasks are similar, he concludes that archival tasks confront librarians with much that is unfamiliar, such as "opening tapes"; making back-up copies; updating originals; reproducing and distributing tapes; providing programs; and, training users in computer techniques; thus, making the merger of the two unworkable.

In "Data Archives as Publishers," (Chapter III) White discusses the most significant distinction between libraries and archives: major archives "publish" data sets. In his discussion, White gives a wealth of detail about the process of publication within data archives--material that is unavailable elsewhere.

Rather than merge libraries and archives White argues for an alternative: using libraries to make information about data more widely available by acquiring from publisher archives the human-readable materials by which the content of data files can be known. Suggestions for better bibliographic control over machine-readable data through traditional library catalogs are given. Because all social scientists are not employed at institutions with data archives, White argues that there is a need for an organization with cross-disciplinary responsibilities--a library or social science information center--to function as a locale for codebook browsing or searching.

Part II of White's thesis, "Buying Patterns at Two Publisher Archives," is
an analysis of the transaction records of the International Data Library and Ref-
erence Service of the University of California, Berkeley (IDL&RS) and the Roper
Public Opinion Research Center in Williamstown, Massachusetts. The analysis is
meant to serve as a basis on which libraries may predict the potential need for
codebooks and data sets. White's intensive exploration of the type of materials
bought (human-readable vs. machine-readable) also presents an interesting pic-
ture of the buyer of the data bases who he hypothesizes tend to be "high-status"
persons (professors and researchers rather than students); from "high-status"
departments (e.g., among the top ten in their fields); male; and geographically
"near" to the archive from which data sets are purchased. After subjecting his
data to statistical tests, White finds that only sex is a predictor and that al-
though those close to IDL&RS buy more, the reverse is true for buying patterns
at Roper.

## Implications for Libraries

The major implication for librarians is that human-readable data sets are valued
by social scientists in their own right (White demonstrates that information about
data is as often purchased as the data themselves) and their purchase could be dele
gated to libraries. The question of the prestige attached to codebook purchases,
as indicated by White's exploration of who buys, is a fascinating insight into the
sociology of the social sciences, and White has brought out a unique factor in ac-
quisitions policy that libraries might well heed.

Reasons for putting human-readable tools, such as codebooks, in academic libraries
include 1) introducing them into institutions that are more numerous and visible
than existing data archives; 2) assuring that the tools are accessible to all dis-
ciplines and professional schools on campus; 3) deploying them with the other bib-
liographic tools and substantive literature available to social scientists; 4)
improving bibliographical information on codebooks by bringing catalogue cards on
them together with cards on associated monographs; and, 5) improving bibliographi-
cal information on data access tools generally by making them discoverable through
the local catalogue, and possibly through union catalogues up to the national lev-
el.

The final chapter of White's dissertation, "Acquisition of Data Access Tools,"
concentrates on guidelines for librarians interested in buying data access tools.
Appendices include correlation matrixes describing buyers at Roper and IDL&RS and
a substantial bibliography on data archives and librarians.

Writing this review from the academic librarian's viewpoint, one is impressed
with the wealth of information marshalled to provide a case for library purchase
of human-readable access tools. If all areas of library buying had as much or-
ganized data as Dr. White provides in this thesis, libraries would move from the
realm of somewhat subjective buying to a more scientific policy. Thus White's
conclusions act at once as a model for acquisitions policy at a universal level,
as well as in the specific case he addresses.

Data archivists, especially those working in isolation from traditional libraries, will want to consider the implications of White's work for their field. Greater acceptance by libraries of access tools to data sets will generate a broader base of potential archive users and, since White argues cogently for the continuing housing of the actual data sets in archives, there is no reason to fear that the very specialized services of archives will be subsumed by libraries without staff expertise to promote their exploitation. We hope that Dr. White will consolidate his groundbreaking findings into articles to be disseminated in the library and data archive press. The thesis itself is a mandatory purchase for library schools and a critical accumulation of data for archivists. We hope that Dr. White's thesis is the first of many dissertations which will explore in greater depth the problems and characteristics of archives and their users.

## QUANTUM COMPLETES SURVEY

Quantum members in Germany have completed a survey of completed, ongoing, and planned research projects in quantitative history. This survey has just been published as Volume I in a new series of Klett Verlag, Stuttgart, the HSF, which will be concerned with quantitative social scientific analysis of historical and process-produced data. The book entitled, The QUANTUM DOCUMENTATION, is available from Ernst Klett Verlag, Rotebühlstrasse 77, Postfach 809, D-7000, Stuttgart 1, at a cost of DM 39. The full bibliographic citation is:

W. Bick, P. J. Müller, H. Reinke. QUANTUM DOCUMENTATION: Quantitative historische Forschung 1977/Quantitative History 1977. Stuttgart: Ernst Klett Verlag, 1977.

## NEW ORGANIZATIONS / REORGANIZATION

## BASS HOSTS CESSDA MEETING TO CREATE IF-DO

Belgian Archives for the Social Sciences Hosts Committee
of European Social Science Data Archives to Create
International Federation of Data Organizations

On the invitation of the Committee of European Social Science Data Archives (CESSDA), representatives of data organizations from the United States, Canada and Europe gathered together at Louvain-la-Neuve, Belgium on the 20th and 21st May, 1977 to discuss and formulate further means of mutual cooperation in the area of social sciences data archiving services.

This meeting was also attended by observers from France, Sweden and Switzerland and a representative of the Division for the International Development of the Social Sciences of UNESCO.

On Sunday, 21st May, they decided on the establishment of an International Federation of Data Organizations (IF-DO), designated Mr. Guido Martinotti (ADPSS, Milan) as its president, and entrusted the duties of secretary to Mr. Erwin Scheuch (ZA, Cologne).

This federation is open to all organizations prepared to participate in the Federation and cooperate in the continuing development of data archiving services in the social sciences.

More information or a copy of the status may be available by contacting the secretary:

> Dr. Erwin Scheuch
> Zentralarchiv für Empirische Sozialforschung
> Universitat zu Köln
> D - 5 Köln (Deutschland)
> Bachermerstrasse, 40

## DATA ORGANIZATION AND MANAGEMENT

The Laboratory for Computer Graphics and Spatial Analysis within the Graduate School of Design at Harvard University has just released a new edition of LAB-LOG, its catalog of computer programs, data bases and publications. Research at the Laboratory is principally concerned with the analysis and graphic display of geographic data used in the planning process. LAB-LOG describes various products which have resulted from this work and which are currently available for distribution to universities, government agencies and private organizations.

LAB-LOG includes a description of six different computer programs for use in the graphical display of spatial data via a line printer, line plotter and cathode ray tubes. A wide variety of cartographic (x-y coordinate) data bases are also described. Publications are available on the subjects of automated cartography, theoretical cartography and theoretical geography. LAB-LOG also contains a brief description of the Laboratory's history, research directions and operating policies

Copies of LAB-LOG are available at a cost of $1.00 each upon request to:

> The Laboratory for Computer Graphics and
>   Spatial Analysis
> 520 Gund Hall - Harvard University
> 48 Quincy Street
> Cambridge, MA 02138

Payment must accompany your order.

## UPCOMING MEETINGS

**IASSIST WESTERN EUROPE**
**june 26 – 29, 1977**

IASSIST West European Meetings

DANSK DATA ARKIV
Hans Christian Andersens Boulevard 38 mezz.
DK-1553 Copenhagen V, Denmark

June 26 through June 29, 1977

PROGRAM

Sunday, June 26, 19.00: (tentative) Welcome and an evening in TIVOLI - concert, ballet, fireworks, etc.

Monday, June 27, 9.30: (At the DDA Offices)

The Issue: What do we need to know about each European Archive? How much detailed information is really required? Can we design one questionnaire that will include all of the required elements and not be a burden to the respondent?

10.30 - 12.30 Small group work

14.00 - 16.30 Small group work

17.00 - 19.00 Written reports from work groups, general evaluation, and plan of action

Tuesday, June 28, 9.30: (At the DDA Offices)

The Issue: Should archive holdings be standardized with respect to documentation? If so, what should be included in these standards? How do we ensure that these standards are adhered to and used by primary research personnel prior to primary analysis?

10.30 - 19.00 (same format as June 27)

Wednesday, June 29, 9.30: (At the DDA Offices)

The Issue: What are the sources of data? Should Archives classify and index these data in a standard manner to facilitate data retrieval as well as cross-national and international research? If so, how should this be accomplished?

10.30 - 19.00 (same format as June 27)

19.00 - 20.00 General summary and plan of action

**QUANTUM**
**august 10 – 12, 1977**

QUANTUM and the Social Science History Association (USA) Announce Joint Conference

"Quantification and Methods in Social Science Research: Possibilities and Problems with the Use of Historical and Process-Produced Data"

University of Cologne

Cologne, August 10-12, 1977

The conference has the following perspectives: presentation of new methodolo ical approaches; discussion of methodological problems in the analysis of various data types; and, discussion of the processing and archiving of historical and process-produced data for research purposes. The emphasis will be on the presentatio of methodological developments for the analysis of particular data types.

Two plenary sessions will deal with general trends and problems of quantitative historical-social science research and the question of international coordination of archiving policies for machine-readable historical and process-produced data.

The preliminary program describes the following sessions: analysis of census data; analysis of life histories; analysis of administrative data; time series analysis; cohort and longitudinal analysis; network analysis; content/document analysis; legislative behavior; thematic cartography; techniques for data transformation/sampling of historical and process-produced data; structure of social bookkeeping-problems with the analysis of complex data files; new data bases; and, preservation, storage and access.

The deadline for registration is 8 July 1977. The conference registration fee is DM 40. Further information can be obtained by writing to QUANTUM e.V., Greinstrasse 2, D-5000, Köln 41, Federal Republic of Germany.

For those IASSIST readers not yet familiar with QUANTUM, further information about the organization (Association for Quantification and Methods in Historical and Sociological Research), activities, and its Newsletter, QUANTUM Information, can be obtained by writing to the address given in the above paragraph.


**EUROPEAN REGIONAL CARTOGRAPHY**
**september 1 – 3, 1977**

Meeting on European Regional Cartography
Bergen 1-3 September 1977

The Norwegian Social Science Data Services has been authorized to organize a meeting to discuss the establishment of a joint data bank for thematic mapping of regional variations in Europe. This will take place at the headquarters in Bergen from 1-3 September 1977.

The Norwegian Research Council for Science and the Humanities has agreed to sponsor the meeting and to cover the local costs. The Norwegian Council has also invited all members of the European Science Foundation to send delegates to the meeting: this request has received the support of the ESF ad hoc Committee for the Social Sciences chaired by Professor Ralf Dahrendorf of the London School of Economics.

Two major social science archives have already agreed to join in the development of the data system and in the production of the first computer atlases: the Belgian Archives in Louvain-la-Neuve and the Archivio dati in Milan. The British SSRC has already agreed to send a representative to the meeting. A group of Swedish social historians has taken steps to organize a file of coordinates for all the administrative units of their country since 1862 and will be invited to present a report and a set of specimen maps at the meeting.

All scholars engaged in the development of facilities for automatic production of thematic maps by region are invited to present reports at the meeting. They are requested to contact the Director of the NSD at Christiesgate 15-19, N-5014 Bergen-Univ, and to submit applications for travel funds to their National Research Councils.

The agenda of the meeting is likely to run as follows:

   I.   Review of development country by country

      1.  Computer facilities for thematic mapping:  technical details
      2.  Registers of localities and of changes in their status and their boundaries over time
      3.  Available files of joint co-ordinates for computer mapping
      4.  Currently used schemes of regionalization level by level:  cf. the scheme set out in European Political Data 19
      5.  Check lists for data resources at different levels of aggregation
      6.  Examples of computer produced maps:  a display will be organized.

   II.  Initial steps towards a link-up of data sets for regional cartography

      1.  The HANDLEY set for the regions of the Six
      2.  The Centre-Periphery Teaching Packages prepared by the Bergen Group:  CH, D, GB, NO
      3.  Projects within the EEC

   III.  Strategies in building a joint European facility

      1.  Mapping systems:  which levels of regionalization should be given priority?
      2.  Data access:  which data sets could be linked up in the first round?  what further steps should be taken?
      3.  First joint products:  what steps could be taken to prepare the first atlases?  which levels of aggregation?  which groups of variables?  descriptive vs analytical atlases?
      4.  Conformant mapping vs point graphics:  alternative procedures for visualizing variations across cities and other densely populated territorial units.
      5.  Establishment of a European Group for Computer Cartography and/or of Data Services for Regional Analysis?

**APDU**
**september 14 – 16, 1977**

The Association of Public Data Users (APDU) announces the theme and preliminary agenda for the 1977 Annual Conference, to be held on September 14-16, 1977, in Washington, D.C.  This year's theme is "Interchange and Acquisition of Public Data."

The agenda includes a review of the year's activities, committee and related reports, planning for 1978, a workshop on selected current survey data products (<u>Current Population Survey</u>, <u>Survey of Income and Education</u>, and <u>Consumer Expenditures Survey</u>), and structured discussions on software interchange and acquisition, interchange of data, and acquisition of data.  Guest speakers at the luncheon sessions are Manuel D. Plotkin, Director of the U.S. Bureau of the Census, and Joseph W. Duncan, Deputy Associate Director, Statistical Policy Division, Office of Management and Budget.

For additional information on the location of the meetings and registration fee, write or telephone Warren G. Glimpse, APDU Annual Conference Chairperson, c/o Data User Services Division, U.S. Bureau of the Census, Washington, D.C. 20233; telephone number:  301/763-1932.


**IASSIST NORTH AMERICA**
**february 8 – 12, 1978**

### Chicago Meeting -- February, 1978

THEME:  State of the Art:  Perspectives

DATE:  February 8-12, 1978

PLACE:  Nordic Hills
        Itasca, Illinois  (limousine service is available from Chicago's O'Hare airport)

PROGRAM COMMITTEE:  Tony Falsetto (Chair), Public Archives of Canada
                    Sheldon Laube, C. M. Leinwond Associates
                    Dick Roistacher, University of Illinois
                    Carolyn Geda, ICPSR
                    Patrick Bova, NORC


WEDNESDAY, February 8

    6:00 p.m. -                         Cocktails, pre-registration

THURSDAY, February 9

| | |
|---|---|
| 8:00 a.m. - 9:00 a.m. | Registration |
| 9:00 a.m. - 10:00 a.m. | Introductions |
| 10:00 a.m. - 10:30 a.m. | Coffee |
| 10:30 a.m. - 12:30 p.m. | Symposia (3) |
| 12:30 p.m. - 2:00 p.m. | Lunch |
| 2:00 p.m. - 4:00 p.m. | Symposia (3) |
| 5:00 p.m. - 7:00 p.m. | Social |

FRIDAY, February 10

| | |
|---|---|
| 9:00 a.m. - 10:00 a.m. | Plenary Session/Introduction to Action Groups |
| 10:00 a.m. - 10:30 a.m. | Coffee |
| 10:30 a.m. - 12:00 p.m. | Action Group meetings |
| 12:00 p.m. - 2:00 p.m. | Lunch/Roundtable discussions |
| 2:00 p.m. - 4:00 p.m. | Action Group meetings |
| 4:00 p.m. - 5:30 p.m. | Plenary Session/Interim reports |
| 7:00 p.m. - | Dinner/Social -- Keynote Speaker |

SATURDAY, February 11

| | |
|---|---|
| 9:00 a.m. - 12:00 p.m. | Action Group meetings |
| 12:00 p.m. - 2:00 p.m. | Lunch/Roundtable discussions |
| 2:00 p.m. - 4:00 p.m. | Action Group meetings |
| 4:00 p.m. - 5:30 p.m. | Plenary Session |
| 5:30 p.m. - 7:00 p.m. | Social |

SUNDAY, February 12

| | |
|---|---|
| 9:00 a.m. - 12:00 p.m. | Final Plenary Session |

SYMPOSIA. The symposia will focus on the Action Group activities and the state of the art. There will be three symposia in the morning and three in the afternoon. Topics for papers and presentations are invited from the general membership. The Action Group "animateur" will chair these sessions; each animateur will be asked to contact his European counterpart and invite papers. Copies of the papers will be distributed in advance if received in time or will be included in the published Proceedings.

ROUNDTABLES. Roundtables will be scheduled for Friday and Saturday. Topics for the roundtables are invited and should be submitted to Tony Falsetto, Public Readable Archives, Public Archives of Canada, 2850 Cedarwood Drive, Apt. 909, Ottawa, Ontario, Canada, with a copy to Carolyn Geda. The roundtables will provide an opportunity to receive feedback, suggestions, and proposals for problems of specific archives and/or discussion of current activities.

## International Sociological Association
### Uppsala -- August 14-21, 1978

THEME: International Interchange of Data

PROGRAM COMMITTEE: John DeVries (Chair), Carleton University, Canada
Elliott Avedon, University of Waterloo, Canada
Don Harrison, National Archives, Washington, D.C., U.S.A.
Guido Martinotti, University of Turin, Italy
Carolyn Geda, Inter-university Consortium for Political and
  Social Research, University of Michigan, U.S.A.
Per Nielsen, Danish Data Archive

THREE SESSIONS (PANELS):

1. Privacy versus Freedom of Information
2. Research Problems Associated with Complex Data Bases
3. Issues in Comparative Data and Research

WORKSHOPS will be scheduled informally.

A BUSINESS MEETING will also be scheduled.

---

### SOCIAL SCIENCE DATA ANALYSIS LABORATORY

#### SEEKS COORDINATOR/LECTURER

To take charge of facilities and services for instructional and re-
search activities including mini-courses, laboratory assistance, and con-
sultation for faculty and students of the Social and Behavioral Science
Departments. Computer programs, data sources and archives, statistical
packages and other computer modelling techniques. Requires experience
with social science data and knowledge of computer languages and file
retrieval methods. Salary $16,000.

Send resume to:      Professor Thomas Muench
                     Dept. of Economics, Room I
                     SUNY at Stony Brook, N.Y.
                     11794

Equal Opportunity Affirmative Action Employer

---

# iassist

## MEMBERSHIP APPLICATION

NAME_____

MAILING ADDRESS_____

_____

INSTITUTIONAL AFFILIATION_____

_____

TELEPHONE NUMBER_____

Date_____          Amount Enclosed_____

Membership fees for calendar year 1977

  Individual: Regular $15 ___          Student $5 ___

  Institutional (two individual memberships): $35 ___

Charter Individual Membership (three years): $100 ___

Institutional Subscription: $25 ___

Payment Enclosed _____ (amount)

       Make check or money order payable to IASSIST.  Send payment to:

                Judith S. Rowe
                PRINCETON UNIVERSITY COMPUTER CENTER
                87 Prospect Avenue
                Princeton, N. J. 08540

For individual memberships, indicate interest in participation in one of the
following action groups:

    DATA ARCHIVE REGISTRY ____          DATA ORGANIZATION AND MANAGEMENT ____

    DATA ARCHIVE DEVELOPMENT ____          CLASSIFICATION ____

    DATA ACQUISITION ____          DOCUMENTATION ____

                         PROCESS-PRODUCED DATA ____

Membership in the IASSIST includes a subscription to the IASSIST Newsletter and a
subscription to S_S Data, F Newsletter of Social Science Technical Acquisitions, each
to appear four times a year. In addition to these subscriptions, the IASSIST member-
ship affords the opportunity to participate at both the international and region-
al levels.

--------------------------------------------------------------------

(for individual's use)

I applied for membership in IASSIST    Fees paid (amount)____  Date_____

# iassist